Forest Management Strategy Model Based on EPIC and Meta-Analysis

Zirun Zhou 1, Chuanqing Sun 2,*, Yujie Xiao 2, Jiahao Zhang 3

1 School of Management Shanghai University, Shanghai, 200444, China
2 School of Management Guangdong University of Technology, Guangzhou, Guangdong, 510520, China
3 School of Economics Shanghai University, Shanghai, 200444, China
* Corresponding author. Email: Edward_Zhang@shu.edu.cn

Abstract. First, to develop a reasonable carbon sequestration model to predict the amount of carbon dioxide sequestered by forests and their products, we constructed an EPIC-based carbon sequestration model to simulate the amount of carbon dioxide sequestered and the actual amount of carbon dioxide sequestered. forest management strategies and find the best strategies. In addition, a meta-analysis-based optimal forest resource utilization model was established, which comprehensively analyzed forests, including carbon dioxide sequestration and other natural, social and economic values.

Keywords: Carbon Sequestration, EPIC, meta-analysis, Forest management strategy.

1. Introduction

To curb climate change, greenhouse gas emissions need to be controlled. In addition to this, the inherent carbon sequestration capacity of the biosphere, especially the carbon sequestration capacity of forests, is required to absorb carbon dioxide.

To develop a reasonable carbon sequestration model to predict the amount of carbon dioxide sequestered by forests and their products, our team considers establishing an EPIC-based carbon sequestration model and bringing in specific environmental parameters to simulate the amount of carbon dioxide sequestered and the actual amount of carbon dioxide sequestered. Different forest management strategies are simulated by changing the parameters to find the best strategy.

To model the optimal use of forest resources, meta-analyses may consider other aspects of forest value, including carbon dioxide sequestration and other natural, human, and economic values. Develop a decision-making model that makes optimal use of the forest. Comprehensively consider the supply of forest fruit products, water conservation, carbon fixation and oxygen release, air purification, soil conservation, nutrient accumulation, forest protection, biodiversity protection, forest recreation, scientific research and education, and other types of services provided by the forest ecosystem.

Next, select several representative forest areas, practically apply the two models that have been established, adjust the duration to 100 years, and then change the corresponding parameters to find the long-term optimal forest management strategy, and get the results within this period. How much carbon dioxide will be sequestered in the selected forest area?

2. Model Establishment and Solution

2.1. Carbon sequestration model based on EPIC model

2.1.1. Model building

This study uses WinEPIC0810Version 6.0 (updated June 2016), a Windows-based version of the EPIC model provided by Texas A&M Agri Life Research Blackland Research & Extension Center. EPIC integrates a tree growth module, hydrological ecology module, nutrient cycle module, and soil erosion module. It comprehensively and quantitatively simulates various physiological and physical
changes of trees in the entire growth cycle with days as the time unit. It is not only a measure of agricultural management strategies and soil and water resources, but also a vital tool for harnessing efficiency.

In EPIC, the yield simulation flow is shown in Figure 1, and the calculation of potential biomass is as follows (Monteith, 1977):

\[
\Delta B_{p,i} = 0.001 \times WA \times PAR_i
\]

\( \Delta B_{p,i} \) represents the potential biomass increase on the day \( i \) (kg ha\(^{-1}\)). WA is the light energy-to-biomass conversion ratio (kg ha\(^{-1}\)/(mg m\(^{-2}\)), and \( PAR_i \) represents the intercepted amount of photosynthetically active radiation (MJ m\(^{-2}\)/d). Calculated using Beer’s law as:

\[
PAR_i = 0.5RA_i(1 - e^{-0.65LAI_i})
\]

RA represents the amount of solar radiation (MJ/m\(^{2}\)), and \( LAI \) represents the leaf area index.

In the presence of environmental stress, the actual daily increase in biomass was calculated by the following method (Williams et al. 1989):

\[
\Delta B_{a,i} = \Delta B_{p,i}S_{reg,i}
\]

\( \Delta B_{a,i} \) is the increased biomass (kg ha\(^{-1}\)), and \( S_{reg} \) represents the minimum stress factor. The model estimates daily environmental availability by inputting daily meteorological days and soil management data and then derives various stress factors by simulating tree demand and environmental availability. If the stress factor is less than 1, it means that a stress phenomenon occurs. At the same time, the model automatically selects the daily minimum stress factor and records the number of stress days for each stress factor year by year through looping nested statements.

Finally, tree yield is calculated using the Harvest Index:

\[
Yield = HIA \sum_{i=1}^{n} \Delta B_{a,i}
\]
**Yield** represents the dry weight of the tree (kg ha\(^{-1}\)), and HIA is the harvest index under the adjustment of water stress, which is calculated by the following formula:

\[
HIA_i = HIA_{i-1} - HI \left(1 - \frac{1}{1 + WSYF \times FHU_i (0.9 - WS_i)}\right)
\]  

(5)

HI as the potential harvest index, WSYF is the tree parameter, indicating the sensitivity of the harvest index to drought, FHU is the tree growth stage factor, and WS is the water stress factor.

EPIC0810 provides five mathematical models for researchers to calculate potential evapotranspiration: Penman-Monteith, Penman, Priestley-Taylor, Hargreaves, Baier-Robertson. Based on the collected meteorological data, this paper uses the Penman-Monteith equation to simulate evapotranspiration (Li Jun et al. 2004a and 2004b).

The forecast of surface runoff is calculated based on the daily precipitation estimation using the SCS curve code equation:

\[
Q = \begin{cases} 
\frac{(R-0.2s)^2}{R+0.8s}, & R > 0.2S \\
0, & R \leq 0.2S 
\end{cases}
\]  

(6)

\(Q\) is the daily runoff (mm), \(R\) is the daily precipitation (mm), and so is the holding force parameter (mm). In different watersheds, the holding force parameter \(S\) also changes due to changes in soil type, land use, slope, and soil moisture content. \(S\) is related to the curve code CN:

\[
S = 254(100/CN_2 - 1)
\]  

(7)

\(CN_2\) is the curve code under humidity state 2 (moderate), which can be found in the SCS Hydrology Handbook. The equation for the subsurface lateral flow is expressed as:

\[
QH_l = (SW_{ol} - FC_l) [1.0 - \exp(-24/TT_{HI})]
\]  

(8)

\(QH_l\) is the lateral flow rate (mm/d) of soil layer \(l\), and \(TT_{HI}\) is the lateral flow transit time (d), which can be estimated by the following formula:

\[
TT_{HI} = TT_l / S
\]  

(9)

\(S\) is the slope of the land surface (m/m).

Moisture vertical penetration:

\[
O_l = (SW_{ol} - FC_l) [1.0 - \exp(-\Delta t / TT_l)]
\]  

(10)

\(O_l\) is the water penetration rate of soil layer \(l\) (mm/d), \(SW_{ol}\) is the soil water content at the beginning of the time interval \(t\) (24h), \(FC_l\) is the soil water holding capacity (mm), and \(TT_l\) is the water penetration rate. The time (h) of soil layer \(l\) can be calculated by the following formula:

\[
TT_l = (PO_l - FC_l) / SC_l
\]  

(11)

\(PO_l\) is the void water content (mm), and \(SC_l\) is the saturated hydraulic conductivity (mm/h), which can be directly input.

Water use refers to the potential water use efficiency from the soil surface to any root depth and is estimated using the following equation:
\[ U_{pi} = \frac{E_{pi}}{1-\exp(-A)} \left[ 1 - \exp \left( \frac{Z}{RZ} \right) \right] \] (12)

\( U_{pi} \) is the total water use rate (mm/d) from the soil surface to the depth \( Z \) (m) on the day \( i \), \( RZ \) is the root zone depth (m), \( A \) is the water use distribution parameter and the root distribution is relatively its value is larger when it is shallow. During tree growth, drought stress may lead to a reduction in tree growth rate, and the stress factor is calculated using the following formula:

\[ WS_l = \sum_{i=1}^{m} \frac{U_{i,l}}{E_{p,i}} \] (13)

\( WS_l \) is the water stress factor, \( U_{i,l} \) is the water use amount in soil layer \( l \), and \( E_{p,i} \) is the potential plant water use amount on the day \( i \). At the same time, the water use rate was used to calculate the effect of water stress on harvest index and tree growth rate:

\[ WUE = 100 \left[ \frac{\sum_{k=1}^{k} U_i - \sum_{k=1}^{k} E_{p,i}}{\sum_{i=1}^{m} U_{i,l} / E_{p,i}} \right] \] (14)

\( WUE \) is the simulated water use rate, \( U_i \) and \( E_{p,i} \) are the actual and potential plant water use rate (mm/d) on the day \( i \), respectively, and \( k \) is the number of days in the growing season.

Nitrogen uptake: A tree's daily N requirement is the difference between the ideal N content and the tree's N content for that day. The amount of N required is estimated using the following formula:

\[ UND_i = (c_{NB})_i B_i - \sum_{k=1}^{k} UN_k \] (15)

\( UND_i \) is the N demand rate of trees (kg/hm\(^2\)), \( (c_{NB})_i \) is the optimal N concentration of trees (kg/t), \( B_i \) is the biomass accumulated on the \( i \)th day (t/hm\(^2\)), and \( UN_k \) is the actual N uptake rate (kg/hm\(^2\)d). The optimal tree N concentration can be calculated as a function of the growth period:

\[ (c_{NB})_i = bm_1 + bm_2 \exp(-bm_3HUL) \] (16)

\( bm_1, bm_2 \) and \( bm_3 \) are the parameters describing the optimal N concentration of trees at the seedling stage, mid-growth stage, and mature stage, respectively. The amount of N supplied by the soil is restricted by the flow of NO\(_3\)-N to the root system:

\[ UNS_{i,l} = \sum_{i=1}^{m} UN_{i,l} \] (17)

\( UNS_l \) is the rate of N supply from soil to plants (kg/m\(^2\)d).

Phosphorus absorption: The calculation method of daily P requirement \( UPD \) and optimal plant P concentration \( c_{PB} \) is similar to the calculation method of N, only need to replace the variables and parameters related to N in equations (17) and (18) with the related terms of P. The soil P supply was calculated using the following equation:

\[ UPS_{i} = 1.5 \cdot UPD_{i} \sum_{l=1}^{m} LF_{ol} \left( RW_l / RWT_l \right) \] (18)

\( UPS_i \) is the amount of P supplied to the soil (kg/hm\(^2\)), \( UPD_i \) is the P requirement of trees (kg/hm\(^2\)), \( LF_{ol} \) is the fast-acting P factor for absorbing P, \( RW \) is the root weight in the upper layer (kg/hm\(^2\)), and \( RWT_i \) is the total root weight on the day \( i \) (kg/hm\(^2\)).
The fast-acting P factor of P absorption is 0.1-1.0, and is calculated by the following formula:

\[ LF_{ol} = 0.1 + \frac{0.9c_{LPl}}{c_{LPl} + \exp(8.01-0.360c_{LPl})} \]  

\( c_{LPl} \) is the available P concentration (g/t) in soil layer \( l \).

The EPIC model simulates dynamic C processes using C routines conceptually similar to the Century model (Figure 4-2). The Century model has been successfully used to simulate soil organic matter (SOM) for various land use types and climates and is one of the models that consistently produces low errors and shows low overall bias. The total carbon pool for soil carbon evaluation comprises multiple units such as a structural layer, metabolic layer, microbial biomass, slow humus, and passive humus. Parton et al. (1987) and Izaurralde et al. describe the initial dynamic C process in the Century model, as well as new C and N modules developed for the EPIC model. These new modules link soil C and N dynamic simulations with tree management, tillage practices, and erosion processes (Izaurralde et al. 2006; Rosenberg et al. 2003). Surface microbial pool turnover was independent of soil texture, which affected the turnover of active SOM (the rate was higher for sandy soils). The model assumes a 60% carbon loss due to microbial respiration at the surface and 55% for all other layers. The allocation of carbon from lignin in structural waste to CO2 was set at 0.3, and the rest was allocated to slow humus. In each process, there are moisture and temperature controls for soil biological processes.

**Figure 2.** The calculation process of the Carbon & Nitrogen model in EPIC

NOTE: XT (temperature control), XW (water control), LMF (fraction of the litter that is metabolic), Lf (fraction of the structural litter that is lignin), Si (fraction of soil mineral component that is silt), CL (fraction of soil mineral component that is clay), Kd (distribution coefficient of organic compounds between soil and liquid in soil), Bd (soil bulk density), theat (soil volumetric water content).

### 2.1.2. Model correction

The coefficient of certainty \( R^2 \), linear root means square error (RMSE), relative linear root means square error (RRMSE), and PBIAS were used to evaluate the simulation accuracy of tree yield and soil moisture content. RMSE can well represent the difference between simulated yield and statistical yield, RRMSE can represent the degree of accuracy of simulation results, and PBIAS reflects the
average trend of SIM greater or less than the OBS value. The smaller the RRMSE, the more sophisticated the simulation, and the closer the PBIAS is to 0, the better the simulation.

![Graph](image_url)

**Figure 3.** Comparison of simulated stored carbon dioxide amount and stored carbon dioxide amount

According to the EPIC model we established, with specific environmental parameters, the simulated amount of carbon dioxide stored and the actual amount of carbon dioxide stored are shown in Figure 3. To sequester more carbon dioxide, it is necessary to strengthen forest logging management, strictly implement logging quotas and wood production plans, and cut down up to 2.37% of the forest area to make wood products. The logging quota and timber production plan are the most basic basis for forest logging.

2.2. A decision model for optimal utilization of forest resources based on a meta-analysis

We will consider other aspects of forest value, including carbon dioxide sequestration and other natural, human, and economic values. Develop a decision-making model that makes optimal use of the forest. Measure the use-value of forests from multiple aspects and find the best management plan for using forests.

The forest ecosystem has various services such as the supply of forest and fruit products, water conservation, carbon fixation and oxygen release, air purification, soil conservation, nutrient accumulation, forest protection, biodiversity protection, forest recreation, scientific research, and education. In the reviewed literature, only a few studies have evaluated the forest protection value and scientific research and educational value of forest ecosystems, so this paper does not include them in the evaluation scope. Here, a meta-analysis of the forest ecosystem service value assessment system is established, and the analysis factors include all the aforementioned types.

The research area of this paper is China (the relevant data of Hong Kong, Macao, and Taiwan in China are lacking, and statistics are not available), and the literature source is China National Knowledge Infrastructure. Since the value assessments in the literature are usually based on different years, to make the data comparable, the values of different assessment base years were adjusted to the price level of 2015 through the Consumer Price Index (CPI). The value after the base year was divided by the forest area in the study area to obtain the unit area value of forest ecosystems in different study areas, which was used as the dependent variable of the meta-regression model.

The corresponding value observations and the mean value of ecosystem services were counted according to the types of ecosystem services, vegetation zoning, and assessment methods (Figure 5). In terms of the number of value observations, the number of value observations for carbon fixation and oxygen release and soil conservation is the largest, and the number of value observations for forest fruit products and nutrient accumulation is relatively small. Most studies use the market value method and shadow engineering method for value evaluation, while the pay willingness method and the travel cost method are used less frequently. From the perspective of ecosystem service value, biodiversity conservation and water conservation provide the highest value, while nutrient accumulation provides the lowest value. The subtropical evergreen broad-leaved forest area has the highest forest ecosystem service value, and no significant difference is found in other areas. Among different assessment
methods, the average value assessed by the willingness-to-pay method is the highest, while the average value assessed by the expense method and carbon tax method is relatively low.

![Graph showing observations and predicted values](image)

**Figure 4.** Observed value predicted value and transfer error (ranked in ascending order of observed value)

In this paper, all continuous variables have been natural logarithmically transformed. Existing studies have shown that the independent variables such as the observed value of ecosystem services and the research area usually have a right-skewed distribution [5], and the logarithmic transformation can effectively reduce the fluctuation degree and heteroscedasticity of the original data, and reduce the impact of high outliers. In a log-log model, the coefficients of continuous variables can be considered elastic coefficients. After necessary processing of various variables, the mean and standard deviation were calculated, and the number of observations was counted.

Our team selects the most widely used type of meta-regression model as the multiple linear regression based on the least-squares method:

$$
\ln y_{ij} = \alpha + \beta^t X^t + \beta^c X^c + \beta^e X^e + \beta^s X^s + u_{ij}
$$

(20)

$y_{ij}$ is the value of forest ecosystem services; $\alpha$ is a constant term; $X$ is an independent variable matrix, where $X^t$ represents the type of ecosystem services, $X^c$ represents the characteristics of the study area (i.e., vegetation zoning and forest area), and $X^e$ represents the surrounding environment (i.e., forest abundance and railway length), $X^s$ represents socioeconomic conditions (i.e., population size and per capita GDP); $\beta$ is the corresponding coefficient matrix; $u_{ij}$ is the random error term.

The underlying assumption of least squares is that there is no correlation between different observations. In the meta-analysis database established in this paper, the maximum number of value observations provided by a document is 8, the minimum is 1, and 96% of the documents provide multiple value observations. Since observations from the same literature are not independent, there may also be correlations between different studies [12], and it is necessary to consider this in model.
building. Some studies use the weighted least squares method with the reciprocal of the number of observations as the weight, which reduces the influence of sample correlation to a certain extent [8].

Weighted least squares corrected for model heteroscedasticity but did not account for dataset hierarchies and differences in randomness between studies. The panel data regression model explicitly considers the heterogeneity between different research objects, weakens the collinearity of variables, and improves the model’s effectiveness. Its specific form is:

\[ \ln y_{ij} = \alpha + \beta^e X^t + \beta^c X^c + \beta^e X^e + \beta^s X^s + w_{ij} \]  

(21)

\( w_{ij} \) is the error term, which can be decomposed into two parts, the former represents the error component of the \( i \)th study, and the latter represents the random observation error. This is an unbalanced panel because the number of valuable observations provided by different studies is not identical. The panel data regression model includes a mixed least-squares model, fixed-effect model, and random effect model. In this paper, the Breusch-Pagan Lagrange multiplier test (LM test) is used to judge the existence of individual random effects, and the Hausman test is used to judge the fixed effect. The pros and cons of effect models and random-effects models.

To choose a model suitable for this dataset, this paper calculates the regression results of the three models, respectively. For the panel data regression model, the results of the LM test showed that there were random effects between different studies (\( P = 0.0839 \)), so the mixed model was not applicable. The results of the Hausman test showed that it was acceptable to have random effects that were independent of the explanatory variables. The null hypothesis (\( P = 0.5859 \)), in which case the random-effects model is preferred. During model building, observations with standardized residuals greater than 1.5 were excluded. Comparing the results of different regression methods, we can see that the overall fitting effect of the panel data regression model is the best. Therefore, this paper uses the random effect model in the panel data regression method to construct the meta-regression equation.

Ecosystem service types, vegetation zoning, forest area, forest abundance, railway lengths, population numbers, and GDP per capita can explain approximately 48% of the value change in total. In the regression results, the regression coefficients of dummy variables (ecosystem service types, vegetation divisions) reflect the direction and degree of deviation of specific variables relative to the control group; the regression coefficients of continuous variables (forest area, forest abundance, etc.) represent the coefficient of elasticity is the ratio of the rate of change of the dependent variable to the independent variable. The specific analysis of the regression results is as follows:

1) Types of ecosystem services: biodiversity conservation, forest fruit products, water conservation, soil conservation, carbon fixation and oxygen release, forest recreation, and air purification. When the conditions were kept constant, the values of the above seven ecosystem service types were all significantly different from the control group (nutrient accumulation), so the value of nutrient accumulation was the lowest. Comparing the regression coefficients, it can be seen that the ecosystem service value of biodiversity protection and water conservation is significantly higher than that of other types, and the ecosystem service value of forest recreation is lower.

2) Vegetation division: Except for the cold temperate coniferous forest area, the regression coefficients of the other six vegetation divisions are all significantly less than 0, indicating that under the condition that other conditions remain unchanged, the forest ecosystem service value of these vegetation divisions is the same as that of There were significant differences in the control group (subtropical evergreen broad-leaved forest area), so the unit area value of the subtropical evergreen broad-leaved forest area was the highest. The regression coefficient of the cold temperate coniferous forest area is positive but not significant, which may be due to the small sample size (only six value observations). Comparing the regression coefficients, it can be seen that the unit area value of the temperate desert area and the alpine vegetation area is significantly lower than that of other types.

3) Forest area: The regression coefficient of forest area is significantly less than 0 (\( P < 0.01 \)), which indicates that the value of forest ecosystem services has a diminishing marginal effect with the increase of area. Under the condition that other conditions remain unchanged, the forest area increases, the total...
value also increases, but the value per unit area decreases. For every 10% increase in forest area, the value per unit area decreases by 1.7%.

4) Forest abundance: The regression coefficient of forest abundance was significantly less than 0 (P < 0.05), indicating that the increase in other surrounding forest areas would result in a decrease in the value per unit area of the study area when other conditions remained unchanged. This may be related to the substitution effect of ecosystem service provision. For every 10% increase in the area of other forests within 50 km, the unit area value of the study area decreased by 0.6%.

5) Railway length: The regression coefficient of railway length is significantly less than 0 (P < 0.1), indicating that under the condition that other conditions remain unchanged, railway construction has a significant negative effect on the ecosystem service value per unit area of the study area. For every 10% increase in railway length within 50 km, the value per unit area will decrease by 0.6%.

6) Population: The regression coefficient of the population is negative, but not significantly different from 0. Continued increases in human populations may degrade ecosystem functions, reducing the value of forest per unit area.

7) GDP per capita: The regression coefficient of GDP per capita is significantly less than 0 (P < 0.05), indicating that under the condition that other conditions remain unchanged, the higher the GDP per capita of the city where the study area is located, the lower the value of forest unit area, Economic development may lead to a decline in ecosystem function. For every 10% increase in per capita GDP, the unit area value decreases by 2.7%.

3. Conclusion

By complementing the two models, the results are more accurate, which has a certain guiding role for practical problems.

Add more influencing factors to the model, including but not limited to the supply of forest fruit products, water conservation, carbon fixation and oxygen release, air purification, soil conservation, nutrient accumulation, forest protection, biodiversity protection, forest recreation, research, and education, etc., so that our model is more accurate.

There is no actual optimal evaluation standard, so our team is committed to finding better evaluation indicators and methods for ecosystem service types to improve our model in future work.

References